

**UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK**

SKILLZ PLATFORM INC.,	:	
a Delaware corporation,	:	
	:	
Plaintiff,	:	
	:	Civil Action No. 1:24-cv-01646-DLC
-against-	:	
	:	
PAPAYA GAMING, LTD., a foreign	:	Hon. Denise L. Cote
corporation; and PAPAYA GAMING, INC., a	:	
Delaware Corporation,	:	
	:	
Defendant.	:	
	:	
	X	
PAPAYA GAMING, LTD.,	:	
	:	
Counterclaim-Plaintiff,	:	
	:	
-against-	:	
	:	
SKILLZ PLATFORM INC.,	:	
	:	
Counterclaim-Defendant.	:	
	X	

**PAPAYA GAMING, LTD.'S AND PAPAYA GAMING, INC.'S MEMORANDUM OF  
LAW IN SUPPORT OF MOTION TO EXCLUDE REPORT AND TESTIMONY OF  
SKILLZ PLATFORM, INC.'S PROPOSED EXPERT DR. ANDREAS GROEHN**

## **TABLE OF CONTENTS**

	<b>Page</b>
TABLE OF AUTHORITIES .....	iii
PRELIMINARY STATEMENT .....	1
FACTUAL BACKGROUND.....	3
A.    Skillz’s Deficient False Advertising Claim .....	3
B.    The Groehn Surveys .....	4
1.    Dr. Groehn’s Target Population and Screener Questionnaire .....	4
2.    The “Perception Survey” .....	5
3.    The “Likelihood Survey” .....	9
I.    LEGAL STANDARDS FOR ADMISSIBILITY OF EXPERT SURVEY EVIDENCE UNDER FEDERAL RULES OF EVIDENCE 702 AND 403 .....	11
II.   THE GROEHN SURVEYS SUFFER FROM NUMEROUS FUNDAMENTAL FLAWS THAT RENDER THEM UNRELIABLE AND IRRELEVANT .....	12
A.    The Surveys Do Not Include Any Controls, Rendering It Impossible to Know Whether or to What Extent the Surveys Are Measuring Anything Relevant .....	13
1.    The Need For Controls Is Firmly Established and Uncontroversial.....	13
2.    Dr. Groehn’s Lone, Novel Defense that Controls Are Impossible And Unnecessary When Testing “Omissions” Lack Any Basis and, In All Events, Only Confirms That His Surveys Are Not Helpful to the Jury .....	15
B.    The Surveys Fail to Replicate Real World Marketplace Conditions .....	17
1.    The Need to Replicate the Real World Is Firmly Established .....	17
2.    Neither Groehn Survey Remotely Replicates Real World Conditions .....	18
3.    Dr. Groehn’s Bald Suggestion That Replicating Real World Conditions “Does Not Matter” When Testing “Omissions” Is Meritless.....	21

C.	The Surveys Are Impermissibly Leading and Highly Biased .....	21
D.	The Surveys Do Not Draw Representative Samples from the Proper Universe .....	23
E.	Even If the Likelihood Survey Were Methodologically Sound, The Results Are Unhelpful and Prejudicial Because, As Dr. Groehn Admits, They Actually Measure Respondents’ Reactions to Being Told They Were Lied To .....	25
F.	The Surveys Suffer From Various Other Methodological Flaws .....	26
CONCLUSION.....		27

**TABLE OF AUTHORITIES****Page(s)****CASES**

<i>Daubert v. Merrell Dow Pharmaceuticals, Inc.</i> , 509 U.S. 579 (1993).....	11
<i>De Lacour v. Colgate-Palmolive Co.</i> , No. 16-CV-8364(KMW), 2024 WL 36820 (S.D.N.Y. Jan. 3, 2024).....	24
<i>Edmondson v. RCI Hospitality Holdings, Inc.</i> , No. 16-CV-2242 (VEC), 2020 WL 1503452 (S.D.N.Y. Mar. 30, 2020) .....	11, 12, 15
<i>J.T. Colby &amp; Co. v. Apple Inc.</i> , No. 11 Civ. 4060(DLC), 2013 WL 1903883 (S.D.N.Y. May 8, 2013), <i>aff'd</i> , 586 F. App'x 8 (2d Cir. 2014) .....	<i>passim</i>
<i>Juicy Couture, Inc. v. L'Oreal USA, Inc.</i> , No. 04 Civ. 7203(DLC), 2006 WL 1012939 (S.D.N.Y. Apr. 19, 2006) .....	1, 15, 17
<i>Kargo Global, Inc. v. Advance Magazine Publishers, Inc.</i> , 06 Civ. 550, 2007 WL 2258688 (S.D.N.Y. Aug. 6, 2007) .....	19
<i>In re KIND LLC "Healthy and All Natural" Litigation</i> , 627 F. Supp. 3d 269 (S.D.N.Y. 2022), <i>aff'd</i> , <i>Bustamante v. KIND, LLC</i> , 100 F.4th 419 (2d Cir. 2024).....	11, 18
<i>MasterCard International Inc. v. First National Bank of Omaha, Inc.</i> , Nos. 02 Civ. 3691(DLC), 03 Civ. 707(DLC), 2004 WL 326708 (S.D.N.Y. Feb. 23, 2004) .....	12, 21
<i>Medisim Ltd. v. BestMed LLC</i> , 861 F. Supp. 2d 158 (S.D.N.Y. 2012).....	12, 15, 17, 23
<i>Proctor &amp; Gamble Co. v. Ultreo, Inc.</i> , 574 F. Supp. 2d 339 (S.D.N.Y. 2008).....	14, 15, 17, 21
<i>Salon FAD v. L'Oreal USA, Inc.</i> , No. 10 Civ. 5063(DLC), 2011 WL 4089902 (S.D.N.Y. Sept. 14, 2011) .....	2, 22
<i>SmithKline Beecham Consumer Healthcare, L.P. v. Johnson &amp; Johnson-Merck Consumer Pharms. Co.</i> , 01 Civ. 2775, 2001 WL 588846 (S.D.N.Y. June 1, 2001).....	16

<i>Souza v. Exotic Island Enterprises, Inc.</i> , 68 F.4th 99 (2d Cir. 2023) .....	15
---	----

## **STATUTES**

Fed. R. Evid. 403 .....	11
Fed. R. Evid. 702 .....	11

Papaya Gaming, Ltd. and Papaya Gaming, Inc. (“**Papaya**”) respectfully submit this Memorandum of Law in Support of their Motion to Exclude the Report and Testimony of the consumer survey expert proffered by Skillz Platform, Inc. (“**Skillz**”), Dr. Andreas Groehn.

### **PRELIMINARY STATEMENT**

The two consumer surveys proffered by Dr. Groehn in support of Skillz’s false advertising claims are fundamentally and fatally deficient on numerous grounds, rendering them both unreliable and unhelpful to the factfinder. Indeed, his “Perception Survey” and “Likelihood Survey” violate multiple fundamental principles of survey design that are firmly established and routinely recognized by this Court as warranting exclusion under Federal Rules of Evidence 702 and 403. Among other issues detailed herein:

1. **Neither survey includes any control** (*infra* § II.A). It is undisputed that the surveys do *not* isolate the impact of the alleged false advertising from consumers’ preexisting beliefs about mobile games or from various forms of survey noise. Dr. Groehn stunningly admitted at deposition that he did not even *attempt* to isolate and measure the alleged false advertising, despite the fact that this is a basic requirement set forth in academic literature about survey design, and Dr. Groehn himself has criticized other experts for the “fatal flaw” of not including controls. As this Court has recognized, the lack of controls requires exclusion of the surveys. *See, e.g., J.T. Colby & Co. v. Apple Inc.*, No. 11 Civ. 4060, 2013 WL 1903883, at \*22 (S.D.N.Y. May 8, 2013) (Cote, J.), *aff’d*, 586 F. App’x 8 (2d Cir. 2014).

2. **Neither survey replicates real-world conditions** (*infra* § II.B). This Court routinely excludes surveys that violate another foundational principle of consumer research by failing to mirror what consumers actually encounter in the marketplace. *See, e.g., J.T. Colby*, 2013 WL 1903883, at \*21 (Cote, J.); *Juicy Couture, Inc. v. L’Oreal USA, Inc.*, No. 04 Civ. 7203, 2006 WL 1012939, at \*25 (S.D.N.Y. Apr. 19, 2006) (Cote, J.). Far from simulating real world

conditions, Dr. Groehn presented respondents with wholly contrived stimuli and scenarios designed to be *unrealistic*. Among other issues, the Perception Survey presented respondents with a *fake* description of a Papaya game that was fabricated by cherry-picking statements from disparate online sources, smashing them together, and stripping them of all their context. Dr. Groehn was forced to admit at deposition that this presentation did *not* reflect an actual “online purchase environment”—a failure that he (correctly) deems deeply problematic when criticizing survey experts in other matters.

3. **Both surveys are extremely leading and biased** (*infra* § II.C). Among other concerns detailed herein, the surveys contain no open-ended or “filter” questions, respondents were primed to focus on bots before answering questions, and respondents were improperly and baselessly told what they already “thought” about their mobile game opponents as part of the premise for questions. The overly suggestive nature of the surveys constitute another ground this Court has relied upon when rejecting survey evidence. *See, e.g., Salon FAD. v. L’Oreal USA, Inc.*, No. 10 Civ. 5063, 2011 WL 4089902, at \*9-10 (S.D.N.Y. Sept. 14, 2011) (Cote, J.).

4. **Both surveys fail to draw representative samples from a proper universe of respondents** (*infra* § II.D). Among multiple other issues: (i) the samples did not include *prospective* Papaya customers despite their relevance and importance to advertising surveys; and (ii) respondents had to report that they played “skill-based” games to qualify for the surveys, but Dr. Groehn recognizes that “skill-based” is an ambiguous “term of art” with a meaning that “depends,” and he has no idea how consumers might have understood that term.

In addition to the foregoing and other deficiencies detailed herein, **the Likelihood Survey is unhelpful to the jury because, as Dr. Groehn admits, the way he framed his question led to results that necessarily encompass *consumer reactions to being told that they***

*were deceived, rather than merely reactions to the presence of bots in mobile games or advertising about bots* (*infra* § II.E). Accordingly, Dr. Groehn is not measuring or isolating the materiality actually pertinent to Skillz’s claims.

Many of the foregoing flaws standing alone are sufficient to warrant exclusion of the surveys. But the *cumulative* effect of all of the flaws—which the Court must consider here—eliminate any doubt that Dr. Groehn’s surveys are woefully deficient, and testimony about them should not be permitted to reach the jury under any circumstances.

### **FACTUAL BACKGROUND<sup>1</sup>**

#### **A. Skillz’s Deficient False Advertising Claim**

The crux of Skillz’s claim is the allegation that Papaya’s advertising deceived consumers with respect to “bots” in Papaya’s mobile games. In its Complaint, Skillz points to language on disparate Papaya “websites” and “application interfaces” that Skillz asserts cause consumers to believe that Papaya games include solely human players—*e.g.*, (1) references to games being “fair” or “skill-based;” (2) a statement that Papaya “has no vested interest in who wins or loses;” and (3) references to matching with “players” or “individuals.” (Complaint (ECF No. 1), at ¶¶ 100-103; 118-121.)

Because Skillz cannot point to any public advertising in which Papaya expressly states there are no bots in games, Skillz’s false advertising theory turns on its ability to prove that the foregoing language “impl[ies]” that games are limited to human players. (*Id.* ¶¶ 111, 113.) Thus, as the Court recognized nearly a year ago: “[R]eferences to ‘players,’ ‘individuals,’

---

<sup>1</sup> Pertinent materials are attached as exhibits to the accompanying Declaration of Jordan A. Feirman (“**Feirman Decl.**”). For efficiency, Dr. Groehn’s April 18, 2025 expert report (excerpts at **Ex. 1**) is referred to and cited as “**Report**,” and excerpts of Dr. Groehn’s May 15, 2025 deposition transcript (**Ex. 2**) is referred to and cited as “**Groehn Tr.**”

‘winners,’ ‘fair,’ and ‘skill-based’ may be found by a jury to *imply* that Papaya’s games . . . are conducted among human players only,” and “the issue” is “whether consumers were *misled*” by Papaya’s advertising “about the presence of bots.” (ECF No. 40, at 7-8 (emphasis added).)

## **B. The Groehn Surveys**

Dr. Groehn opines about two surveys he designed and fielded in April 2025: a “Perception Survey” and a “Likelihood Survey.”

### **1. *Dr. Groehn’s Target Population and Screener Questionnaire***

The “target population” for Dr. Groehn’s surveys was “those who have used mobile gaming apps during the relevant time period”—which “time period” Dr. Groehn understood to mean “the past four years.” (Report ¶¶ 31-32, 56.) Both surveys thus used the same criteria to qualify respondents, including that respondents had to self-report they “played skill-based mobile games in the past four years (since 2021).” (Report ¶¶ 22-24, 51; Feirman Decl. Ex. 3 (Perception Survey Questionnaire) at 5, Q.13.) Prospective game players who had not already played such games thus could not qualify for either survey.

Neither survey defined the term “skill-based.” (Groehn Tr. at 212:8-12.) Dr. Groehn testified, however, that “skill-based” is a “term of art” and its meaning in the mobile gaming context “depends;” as a result, he “do[es]n’t know” how respondents understood that term in the survey or whether they had the same understanding as him.<sup>2</sup> (*Id.* at 205:3-12; 207:4-5; 212:13-213:15; 214:14-215:14; 222:7-21.)

---

<sup>2</sup> Evidencing the likely confusion among respondents concerning what the survey meant by “skill-based” gaming, nearly 1,200 people failed to qualify for the surveys because they reported they had not played a “skill-based” game, despite having already responded to a previous question that they played a “qualifying” mobile game like Solitaire or Blackjack. (Report ¶¶ 44, 65, Tables 1& 3.)

In designing his surveys, Dr. Groehn did not consider any information concerning the demographic characteristics of Papaya’s target or actual consumer base, nor did he set quotas to attempt to match his samples to those characteristics. (Groehn Tr. at 343:6-10.) Instead, his survey vendor sent out invitations to potential survey takers “according to current Census data” (Report ¶ 35), and Dr. Groehn asserts that the resulting distribution of qualifying respondents people must accurately reflect Papaya consumer demographics (Groehn Tr. at 337:18-339:2).

## 2. *The “Perception Survey”*

### (a) *The Perception Survey’s Purpose to Measure Whether Papaya Advertising Caused Consumer Deception*

The expressly stated “purpose” of the Perception Survey “is to determine whether consumers interpret Papaya’s advertising as conveying whether they will play against only humans, only computerized bots, or a mix of both humans and bots in tournaments/matches on Papaya’s platform.” (Report ¶ 18.) Dr. Groehn recognizes that he was not seeking “to understand consumers’ general attitudes or beliefs about Papaya’s platform and games” (Groehn Tr. at 121:3-15), but instead to measure whether alleged confusion about the nature of opponents in Papaya’s games is specifically attributable to Papaya’s advertising.

Yet, at deposition, Dr. Groehn newly indicated that his survey did *not* measure the impact of any statements by Papaya that Skillz has challenged as deceptive (*e.g.*, “fair,” “skill-based,” “players”). Rather, he repeatedly insisted that *his survey only was seeking to measure the impact of a purported “omission” of bot disclosures in Papaya advertising*—*i.e.*, that such “omission” was “the source of consumer deception” and “caused consumers to be deceived” (*id.* at 121:16-122:10; 127:7-12; 130:7-11; 141:2-6; 144:15-20; 177:15-19; 255:6-256:6):

Q: [Y]our Perception Survey is seeking to measure whether something allegedly deceptive in Papaya’s advertising is the source of consumer deception; right?

A: The omissions by Papaya of the fact that they were using bots, we’re measuring that, yes.

\* \* \* \*

Q: So your survey was only testing the impact of an omission in Papaya advertising, not the impact of any affirmative statements in Papaya advertising?

A: Yes.

\* \* \* \*

Q: Are you measuring the confusion caused by anything else other than the omission of language that says bots are present in the game?

A: I don’t think so.

(b) Dr. Groehn’s Deeply Flawed Methodology and Stimuli

After qualifying, respondents were told they “will be shown a short video advertisement related to a mobile gaming app where you wager your money to play against other players and the winner of the game is awarded cash or a prize.” (Feirman Decl. Ex. 3 (Questionnaire) at 7.) On the next screen, respondents clicked a button to start the video, but a statement immediately above the video previewed an upcoming attention-check question by noting: “You will be asked to enter the number you see at the end of the video below.” (*Id.* Ex. 4 (Screenshots) at 3.)<sup>3</sup> Dr. Groehn thus conceded at deposition he had directed respondents in advance to pay particular attention to the final seconds of the video. (Groehn Tr. at 364:18-365:8.)

Respondents were shown a thirty-second video ad for Papaya’s “Solitaire Cash” game, which did *not* include the advertising language challenged by Skillz in this case. (Report ¶ 25.) Dr. Groehn did not (and still does not) know any information—nor does he recall ever asking—

---

<sup>3</sup> Despite having designed the study, Dr. Groehn disconcertingly does not know whether the video could be paused or viewed again by respondents. (*Id.* at 360:8-14.)

about: where or how often the ad was shown; how many people were exposed to it; what period of time it ran; what media it ran on; or whether it is “representative of the content of other marketing materials Papaya used.” (Groehn Tr. at 226:7-228:2.) Apart from one other video he did not list in his Report and about which he could recall nothing at deposition, Dr. Groehn did not consider using any other piece of actual Papaya advertising as a survey stimulus. (*Id.* at 19:19-22:16; 228:3-13.)

After viewing the video, respondents were asked a single closed-ended question whether, “[r]ecalling the video advertisement,” they thought the “other players in that game would be” “only human players,” “only computerized players/bots,” or a “mix” (Report ¶ 26, Fig. 6):

Recalling the video advertisement you just watched, if you were going to play the game advertised in the video, do you think that the other players in that game would be:

- ☐ A mix of both human players and computerized players/bots
- ☐ Only computerized players/bots
- ☐ Only human players
- ☐ Not enough information
- ☐ Don't know/Not Sure

Although this question clearly primed respondents to focus on the issue of bots, they were nevertheless shown another stimulus and asked the same question again. Respondents were instructed to “imagine that after viewing the advertisement, you click on the embedded link, which takes you to the app store screen for the game that was advertised.” Respondents were then told to review “the description of the game below,” after which they were asked the same closed-ended question about whether they thought the game included only humans, only bots, or a mix. (*Id.* ¶¶ 27-28, Fig. 7.) The “description” was four lines of plain text shown as follows:

“Solitaire cash makes sure to match players against other opponents with a similar skill level to ensure a fun and fair experience for everyone. You’ll be matched with other players within the same skill level, and you all will get the same deck —so the game is totally fair and skill-based. The outcome of Solitaire cash is based on the skill of the players, rather than luck or chance. Solitaire cash has no vested interest in who wins or loses, nor does it profit on the outcome of a tournament that we provide.”

Dr. Groehn refers to this as a “mock app store description” because *he invented the*

*description solely for the survey.* (*Id.*) In reality, it is a compilation of language challenged in Skillz’s Complaint, cherry-picked from four different online sources—one of which was a Canadian webpage not targeted at U.S. consumers—and “stripped” of all of their surrounding context. (Report ¶ 27 n.9; Groehn Tr. at 239:5-18; 244:10-13; 247:1-9; 274:18-276:17.)<sup>4</sup>

Although Dr. Groehn tellingly and repeatedly refused to answer the simple question at deposition (*id.* at 239:19-244:9; 270:6-273:11), these four sentences never appear together in the real world marketplace. Even where snippets of the language appear on their own in individual FAQ pages or app listings, they are only viewable under specific circumstances (*e.g.*, after scrolling or clicking to expand a previously hidden section of a webpage), and appear among a great deal of other text and imagery that Dr. Groehn removed. (*Id.* at 273:22-289:3; Feirman Decl. Exs. 5-8.) Dr. Groehn thus conceded that (1) he removed content “that actually reflect[s] an online purchase environment” (Groehn Tr. at 244:14-18), and (2) he has no understanding of whether consumers in the real world actually attend to the language he presented via the fake stimulus (*id.* at 247:10-20).

All respondents in the Perception Survey were shown the same, non-randomized sequence of two stimuli and the follow-up question, and nothing more. There were no controls. (*Id.* at 154:13-16.)

### (c) Reported Results

Because his Perception Survey had no controls, Dr. Groehn’s results are only reported as gross percentages. Following the video, 45.8% of respondents thought the game would include

---

<sup>4</sup> Dr. Groehn testified that, even with respect to this “mock” stimulus that he admits “focus[es] on the four statements that Skillz alleges are deceptive,” he was still only “seeking to determine the impact of an omission in Papaya’s advertising.” (*Id.* at 144:15-20; 247:1-9.) That confusing testimony can only be understood as a function of Dr. Groehn’s after-the-fact, but meritless, attempt to defend his utter lack of any controls (discussed *infra* at § II.A).

“only human players,” while a slightly higher percentage thought the game would include a mix of humans and bots. (Report ¶ 47, Fig. 10.) When asked the question again following presentation of the “mock” stimulus, and after having been primed to focus on the issue of bots, 60% of respondents thought the game would include only humans. (*Id.* ¶ 48, Fig. 11.)

Dr. Groehn concludes that his results show “significant customer confusion” (*id.* ¶ 72) and that “Papaya’s advertising materials have the potential to deceive a substantial portion of the target audience into believing that Papaya games only include human players” (*id.* ¶ 49).

### 3. *The “Likelihood Survey”*

Dr. Groehn purports to draw conclusions from his oddly-named “Likelihood Survey” about what “is material to the target audience” for mobile gaming apps. (Report ¶ 70; Groehn Tr. at 98:11-99:9.) The expressly stated purpose of the Likelihood Survey was “to determine how likely consumers of mobile gaming apps would be to stop or continue playing and spending money on a mobile gaming app if they were informed that some of their opponents in the game whom they thought were real humans were actually bots.” (Report ¶ 50.)

The Likelihood Survey did not include any stimulus or any control. (*Id.* ¶¶ 52-53; Groehn Tr. at 154:13-16; 299:6-22.) Respondents were told to “[c]onsider your recent experience playing a skill-based, mobile app game where you paid an entry fee for the opportunity to win cash or prizes.” (Report ¶ 53 & Fig. 12.) “Skill-based” was not defined, and “recent” could have meant as far back as four years prior. (Groehn Tr. at 218:12-219:16.)

Whatever “recent experience” with an unspecified app respondents may have been considering—which Dr. Groehn conceded could be “any skill-based game offered by any platform” (*id.* at 222:1-6)—they were asked only one closed-ended question pertinent to Skillz’s claims: how “likely would you be to” either “stop” or “continue” “playing and spending your money on that app” specifically “[if] you were informed that some of the other players in your

game who you thought were real people were actually bots” (Report ¶ 53 & Fig. 12):<sup>5</sup>

16. If you were informed that some of the other players in your game who you thought were real people were actually bots, how likely would you be to stop playing and spending your money on that app?

- ☐ Very unlikely
- ☐ Unlikely
- ☐ Neither likely nor unlikely
- ☐ Likely
- ☐ Very likely
- ☐ I don't know / not applicable

Even though he told respondents what they “thought,” Dr. Groehn acknowledged that he does not “know whether any of the respondents . . . had any particular thoughts about whether other players in their game were humans or bots.” (Groehn Tr. at 309:19-310:1.) The survey provided no context concerning the frequency, nature, or operation of the hypothetical bots that could have impacted respondents’ reactions to the app they had in mind. (*Id.* at 328:19-332:4.)

Among respondents asked about their likelihood to “continue” “playing and spending money” upon being “informed” players they “thought were real people” were “actually bots,” 51.2% were “unlikely or very unlikely” to do so, while 34% were “very likely” or “likely” to do so. (Report ¶ 68, Table 5.) Among respondents asked about their likelihood to “stop” playing and spending upon being so “informed,” 66.8% were “likely” or “very likely” to do so, while 17.2% were “unlikely or very unlikely” to do so. (*Id.*, Table 6.) Based on these figures, Dr. Groehn concludes that “appropriate disclosure as to the presence of bots in a game is material to the target audience and directly related to whether consumers will keep playing.” (*Id.* ¶ 70.)

<sup>5</sup> The survey also included a question about likelihood to continue or stop if respondents were informed they “would have to wait one hour to learn about the results of [their] game.” But the Groehn Report nowhere states why that question was asked, and Dr. Groehn’s deposition testimony provided no clarity about the question’s relevance. (Groehn Tr. at 300:11-304:9.)

## ARGUMENT

### **I. LEGAL STANDARDS FOR ADMISSIBILITY OF EXPERT SURVEY EVIDENCE UNDER FEDERAL RULES OF EVIDENCE 702 AND 403**

Federal Rule of Evidence 702 provides that:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if the proponent demonstrates to the court that it is more likely than not that: (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue; (b) the testimony is based on sufficient facts or data; (c) the testimony is the product of reliable principles and methods; and (d) the expert’s opinion reflects a reliable application of the principles and methods to the facts of the case.

Fed. R. Evid. 702. The court acts as a “gatekeeper,” with “responsibility to ensure that the expert is qualified to give such testimony and that the expert’s opinions emanate from reliable methods, reliably applied.” *J.T. Colby*, 2013 WL 1903883, at \*19 (Cote, J.); *see Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 592-93 (1993).

“The Second Circuit has distilled Rule 702’s requirements into three broad criteria:

(1) qualifications; (2) reliability; and (3) relevance and assistance to the trier of fact.” *In re KIND LLC “Healthy and All Natural” Litig.*, 627 F. Supp.3d 269, 282 (S.D.N.Y. 2022), *aff’d*, *Bustamante v. KIND, LLC*, 100 F.4th 419 (2d Cir. 2024). Accordingly, the proffering party must demonstrate, by a preponderance of evidence, that “(1) the expert is qualified; (2) the proposed opinion is based on reliable data and methodology; and (3) the proposed testimony would be helpful to the trier of fact.” *Edmondson v. RCI Hosp. Holdings, Inc.*, No. 16-CV-2242, 2020 WL 1503452, at \*3 (S.D.N.Y. Mar. 30, 2020).

The Court also analyzes admissibility of expert evidence under Rule 403, which permits exclusion “if its probative value is substantially outweighed by . . . unfair prejudice, confusing the issues, [or] misleading the jury.” Fed. R. Evid. 403. “Because of the danger that expert evidence will mislead the jury, a court weighing admissibility under Rule 403 exercises more

control over experts than lay witnesses.” *MasterCard Int’l Inc. v. First Nat’l Bank of Omaha, Inc.*, No. 02 Civ. 3691, 2004 WL 326708, at \*7 (S.D.N.Y. Feb. 23, 2004) (Cote, J.).

Rules 702 and 403 are frequently applied by this Court to exclude proposed survey evidence. *See, e.g., J.T. Colby*, 2013 WL 1903883, at \*20-23 (Cote, J.) (excluding survey on Rule 702 and 403 grounds); *MasterCard Int’l*, 2004 WL 326708, at \*7-10 (Cote, J.) (same). “[A] survey should be excluded under Rule 702 when it is invalid or unreliable, and/or under Rule 403 when it is likely to be insufficiently probative, unfairly prejudicial, misleading, confusing, or a waste of time.” *Medisim Ltd. v. BestMed LLC*, 861 F. Supp. 2d 158, 167 (S.D.N.Y. 2012). While individual flaws may independently warrant exclusion, Rules 702 and 403 further “require the court to *look at the cumulative effect of all of the flaws in a survey* in determining its admissibility.” *Edmondson*, 2020 WL 1503452, at \*6 (citations omitted; emphasis added).

To be admissible, a survey must be “properly designed, executed, and described,” and “conducted in accordance with generally accepted survey principles.” *Id.* (quoting Shari Seidman Diamond, *Reference Guide on Survey Research*, in *Reference Manual on Scientific Evidence*, 359, 362-64 (Federal Judicial Center, 3d ed. 2011)). In assessing reliability, courts consider factors including whether, in pertinent part: (1) the proper universe was examined and a representative sample was drawn from that universe; (2) methodology and execution complied with generally accepted standards in the field of such surveys; (3) questions were leading or suggestive; and (4) questions were clear or ambiguous. *Id.*; *Medisim*, 861 F. Supp. 2d at 166.

## **II. THE GROEHN SURVEYS SUFFER FROM NUMEROUS FUNDAMENTAL FLAWS THAT RENDER THEM UNRELIABLE AND IRRELEVANT**

Both the Perception Survey and Likelihood Survey are rife with serious methodological

and analytical deficiencies.<sup>6</sup> Many of these deficiencies are so fundamentally problematic that they would independently suffice to require the surveys' exclusion. The cumulative effect of these flaws, however, eliminates any doubt that Skillz cannot possibly carry its burden to demonstrate that these surveys should reach a jury.

**A. The Surveys Do Not Include Any Controls, Rendering It Impossible to Know Whether or to What Extent the Surveys Are Measuring Anything Relevant**

**1. The Need For Controls Is Firmly Established and Uncontroversial**

Controls are essential in false advertising surveys. Without them, there is no way to ensure that the survey results are isolating and measuring the impact of a party's alleged false advertising (*i.e.*, the impact of the conduct challenged as unlawful) rather than just capturing consumers' preexisting beliefs about a product, or various forms of survey "noise."

Academic literature on survey design for Lanham Act litigation, including sources expressly cited and relied upon by Dr. Groehn in his Report, universally recognize the need for controls. For example, the oft-cited *Reference Guide on Survey Research* from Professor Shari Seidman Diamond (hereinafter, "**Reference Guide**") provides a salient example demonstrating why controls are necessary when seeking to evaluate (among other things) how "the content of a commercial affect respondents' perceptions or understanding of a product or commercial:"

[R]espondents assigned to the experimental condition view an allegedly deceptive commercial, and respondents assigned to the control condition either view a commercial that does not contain the allegedly deceptive material or do not view any commercial. Respondents in both the experimental and control groups answer the same set of questions about the allegedly deceptive message. . . . If 40% of the respondents in the experimental group responded indicating that they received the deceptive message . . . whereas only 8% of the respondents in the control group gave that

---

<sup>6</sup> Papaya is not challenging Dr. Groehn's qualifications to opine about surveys generally. That said, his credentials and experience reveal that his true expertise lies in economics and not false advertising surveys (Report ¶¶ 2-4; Groehn Tr. at 63:11-64:6; 99:21-103:15), which may help explain why his methodology is so flawed and why his defenses of bizarre survey design choices (*e.g.*, not using any control, relying on extremely doctored stimuli) find no support in any academic literature or caselaw.

response, the difference . . . can be attributed only to the allegedly deceptive message. ***Without the control group, it is not possible to determine how much of the 40% is attributable to respondents' preexisting beliefs or other background noise[.]***

(Feirman Decl. Ex. 9, at 397-401 (emphasis added.)) Similarly, the ABA Treatise on “Trademark and Deceptive Advertising Surveys”—also cited and relied upon by Dr. Groehn in his Report—emphasizes the need for controls, including in a chapter devoted specifically to false advertising surveys. (Feirman Decl. Ex. 10, at 212-19 (controls address “noise” and “bias” such as “existing consumers’ preconceptions, bias, or reactions to anything other than the test stimulus . . . by isolating the challenged claim from other possible effects”).

These basic principles are uncontroversial. Dr. Groehn himself has criticized other Lanham Act surveys for the “fatal” failure to include a control, including twice recently in a case that went to trial in the Southern District of California, *Iprimisrx, LLC v. OSRX, Inc.*:

- Dr. Groehn criticized an advertising survey for failure to include a control, quoting from the Reference Guide that courts “more often recognize the absence of a control as a fatal weakness,” and noting that “[w]ithout a control group there is no way to determine whether the observed effects are due to some baseline level of error that is unrelated to the proposition being tested.” (Feirman Decl. Ex. 11, ¶¶ 25-28).
- Late last year, Dr. Groehn testified before a jury that the lack of a control in a secondary meaning survey was “in itself a fatal flaw in [the expert’s] methodology.” (Feirman Decl. Ex. 12, at 492:19-495:15.)

This Court, too, has stated that “[c]ontrols are an essential feature of reliable survey evidence because they enable the surveyor to separate the wheat (the effect of the advertisement, alone, on the participant) from the chaff (the effect of the participant’s prior knowledge and/or prior (mis)conceptions).” *P&G v. Ultreo, Inc.*, 574 F. Supp. 2d 339, 351 (S.D.N.Y. 2008) (citations omitted); *see also* Feirman Decl. Ex. 10 (ABA Treatise) at 214 (“Today, most courts will reject or give little weight to surveys with no control group to control for consumers’ preexisting beliefs.”) This Court has thus routinely recognized the failure to include a control as

a fatal deficiency requiring exclusion of Lanham Act surveys. *See, e.g., Souza v. Exotic Island Enters., Inc.*, 68 F.4th 99, 113 (2d Cir. 2023) (affirming exclusion where this court “flagged the study’s absence of a control group”); *Edmondson*, 2020 WL 1503452, at \*7 (noting the “failure to use a control group also makes it difficult to . . . account for participants’ prior beliefs, biases, guessing, and other external factors”); *J.T. Colby*, 2013 WL 1903883, at \*22 (Cote, J.) (“In order to offer sound results, most surveys must employ an adequate control”); *Medisim*, 861 F. Supp. 2d at 180 (failures with respect to choice of control were “deeply troubling, and indicative of a serious flaw in the design”); *Ultreo*, 574 F. Supp. 2d at 352 (rejecting false advertising survey, noting that “without a control, one cannot rely on the results of the study for any purpose”); *Juicy Couture*, 2006 WL 1012939, at \*26 (Cote, J.) (finding it “particularly remarkable” that survey failed to include a control).

**2. Dr. Groehn’s Lone, Novel Defense that Controls Are Impossible And Unnecessary When Testing “Omissions” Lack Any Basis and, In All Events, Only Confirms That His Surveys Are Not Helpful to the Jury**

Despite the fact that Dr. Groehn sought to measure whether Papaya’s advertising caused consumer deception and changes in consumer behavior, his surveys contain no controls, and his Report does not even mention them. Remarkably, Dr. Groehn testified that *he did not even “attempt[ ] to isolate the impact of [the] alleged omission in Papaya’s advertising” from preexisting consumer beliefs or survey noise.* (Groehn Tr. at 154:17-155:6.)

Instead, Dr. Groehn raised at deposition an entirely new defense for his violation of basic survey principles, arguing that a control was not possible or necessary because he was testing an “omission” from advertising, and “[a] control does not work with an omission.” (*Id.*; *see also id.* at 157:8-18; 173:13-17.) Dr. Groehn could not identify a single authority or judicial decision

supporting his theory,<sup>77</sup> nor does any of the literature cited in his Report suggest an exception from controls when testing “omissions;” rather, Dr. Groehn could only baldly and self-servingly claim that it is “obvious” and “common knowledge.” (*Id.* at 155:7-18; 157:19-158:12; 160:13-164:14; 165:6-168:17; 178:14-181:20.)

That Dr. Groehn relies exclusively on his own say-so is not surprising, because his theory makes no sense. Skillz is well aware of this; the theory was roundly rejected by Papaya’s survey expert, Dr. Bruce Isaacson, when asked about it by Skillz’s counsel at deposition. (Feirman Decl. Ex. 13 at 221:9-222:15.) Dr. Isaacson, who has decades of advertising survey experience, explained that testing omissions with controls is “quite common” and identified examples of surveys that were testing omissions. (*Id.* at 160:9-168:13; 225:8-227:16.)

Regardless, even if controls somehow did “not work with omissions,” that would not change the fact that Dr. Groehn’s results are not helpful to the factfinder. Because Dr. Groehn admits that his surveys fail to isolate the impact of the alleged false advertising on consumers, his gross percentage results are a black box. Just as in Professor Diamond’s example above, it is not possible to determine how much of Dr. Groehn’s results are attributable to respondents’ preexisting beliefs or survey noise, and thus would not be indicative of deception or behavior actually caused by Papaya’s alleged “omission” or unlawful conduct.

Accordingly, if jurors hear, for example, that 45.8% of respondents thought Papaya games included “only humans” after seeing a “mock” description, or that 51.2% were “unlikely” to continue “playing and spending their money” in a hypothetical scenario, it is (at bare

---

<sup>77</sup> By contrast, this Court has excluded surveys that lack controls even when the alleged false advertising concerns “omissions.” *See, e.g., SmithKline Beecham Consumer Healthcare, L.P. v. Johnson & Johnson-Merck Consumer Pharms. Co.*, 01 Civ. 2775, 2001 WL 588846, at \*2, 11-13 (S.D.N.Y. June 1, 2001) (testing commercial that “obscured” portions of a bar graph regarding a heartburn relief study, thus “omitting, if not affirmatively concealing” certain results of that study).

minimum) confusing and prejudicial because the jurors cannot understand whether any of those respondents were allegedly deceived or influenced by Papaya’s advertising as opposed to responding based on their preconceived notions about mobile games or a confound in the survey design. This failure alone warrants exclusion of the surveys. *See, e.g., J.T. Colby*, 2013 WL 1903883, at \*22 (Cote, J.) (excluding survey for lack of control); *Juicy Couture*, 2006 WL 1012939, at \*26 (Cote, J.) (same).

**B. The Surveys Fail to Replicate Real World Marketplace Conditions**

**1. *The Need to Replicate the Real World Is Firmly Established***

As this Court has recognized, a Lanham Act survey “should make some attempt to replicate marketplace conditions” because “the Lanham Act does not protect against confusion in the abstract, instead, it protects consumers from confusion in the marketplace.” *J.T. Colby*, 2013 WL 1903883, at \*20 (Cote, J.). Thus, “[t]he failure of a survey to approximate actual marketplace conditions can provide grounds for inadmissibility.” *Medisim*, 861 F. Supp. 2d at 166; *see Juicy Couture*, 2006 WL 1012939, at \*25 (Cote, J.) (finding survey has “no value” because of “fundamental flaws in its design,” including “[m]ost significantly, the survey did not replicate the marketplace conditions in which consumers encounter [the relevant] products”).

If a survey tests only stimuli that are materially different from, or not fairly representative of, advertising to which actual consumers are exposed in the real world, the probative value of the survey is highly dubious. After all, if respondents are seeing significantly altered substantive messages or context from what actual consumers would be seeing, those respondents’ perceptions cannot reasonably be extrapolated to what is happening in the marketplace. Similarly, if a survey only tests stimuli to which few consumers are exposed in the real world, whatever application survey responses may have to real world conditions is minimal, at best. *See, e.g., Ultreo*, 574 F. Supp. 2d at 352 (court could not conclude “how consumers would react

to the actual live ad” where survey showed respondents “materials that were substantially altered, as well as advertising materials that are no longer being used”).

These principles, too, are not controversial, and Dr. Groehn himself has criticized other survey experts for failures to replicate real world conditions. For example, he recently opined that survey respondents should be “presented with an experience that mirrors the real world as closely as possible,” criticizing a Lanham Act survey for “mak[ing] no attempt to replicate an online marketplace environment” when the stimuli were presented “without any graphics or images that reflect an online purchase environment.”<sup>8</sup> (Feirman Decl. Ex. 14, ¶¶ 29-32.) Dr. Isaacson agrees, and has opined in this case that reasonably simulating marketplace conditions is a critical and generally accepted principle of false advertising surveys. (*See, e.g., id.* Ex. 13 at 47:23-48:13; 94:9-16; 222:11-15.)

## **2. *Neither Groehn Survey Remotely Replicates Real World Conditions***

The Perception Survey fails to replicate real world conditions in at least two ways:

**First**, respondents were presented with the “mock app store description” that appears nowhere in the real world. As detailed *supra*, this “Franken-stimulus” was comprised of statements challenged by Skillz that were stitched together from different sources (including sources directed at *Canadian* consumers) and “stripped” of all context. Dr. Groehn admits both that he: (1) “provided no graphics or images that actually reflect an online purchase environment;” and (2) he did not know (or care) whether consumers in the real world actually attended to the statements, even though in the real world the statements are placed among other

---

<sup>8</sup> At deposition, Dr. Groehn suggested that the requirement to mirror real world conditions was limited to a trademark confusion survey. (Groehn Tr. at 266:11-268:14.) There is no principled basis, let alone authority that holds, that this generally accepted principle does not apply to false advertising studies.

text and imagery and very well might not be noticed. Accordingly, there can be no genuine dispute that the “mock” stimulus is an utterly unrealistic depiction of Papaya advertising and manifestly improper for evaluating the impact of such advertising (whether arising from an alleged “omission” or otherwise). *See, e.g., In re Kind*, 627 F. Supp. 3d at 288 (rejecting “mock-up” stimulus showing alleged deceptive ad claim “in isolation, rather than as part” of a statement as it actually appeared on products); *J.T. Colby*, 2013 WL 1903883, at \*21 (Cote, J.) (excluding surveys that “made no attempt to replicate market conditions and deprived the survey takers of every contextual clue they would encounter” when looking at products); *Kargo Glob., Inc. v. Advance Mag. Publ’rs, Inc.*, 06 Civ. 550, 2007 WL 2258688, at \*11 (S.D.N.Y. Aug. 6, 2007) (noting “improper” use of stimulus that was “not an actual advertisement” and “would never be encountered by . . . prospective end users”).

**Second**, Dr. Groehn used a Papaya video as a stimulus but he lacked (and did not seek) any information whatsoever concerning that video. As noted *supra*, Dr. Groehn knows nothing about where or how often the ad was shown; how many people were exposed to it; what period of time it ran; or on what platforms it ran. Critically, Dr. Groehn did not even bother to review other Papaya advertising to determine whether the video was “representative of the content of other marketing materials Papaya used.” (Groehn Tr. at 227:11-14.) As a result, there is no basis to presume that the video is an appropriate exemplar of Papaya advertising to which a significant portion of Papaya consumers would be exposed in the real world, and respondents’ interpretation of that video has extremely limited (if any) real-world applications.

The Likelihood Survey is wholly unrealistic in multiple ways as well. For example, as detailed *infra* at § II.D, the survey purports to tell respondents what they already “thought” about other app players, telling respondents that they were misled about the presence of bots. But Dr.

Groehn had no basis to presume—let alone data to establish—that respondents actually had those preexisting “thoughts” about other players.<sup>9</sup> The survey thus sets up an unsupported hypothetical that cannot be confirmed to replicate a real world scenario.

The Likelihood Survey also fails to mirror real world conditions because the question asking respondents how their behavior might change if they were informed about bots (or that they were deceived) provided no additional context that would reasonably inform respondents’ reactions. For example, there are no references to any of numerous factors concerning potential bot use—including specifically alleged uses of bots by Papaya—that could have mattered significantly to whether a consumer would continue or stop playing a mobile game—*e.g.*, if bots made the respondent more likely to win, if the bots were matched against players based on skill-level, the relative frequency of the bots, etc. Dr. Groehn conceded at deposition that such additional context could matter (Groehn Tr. at 327:18-333:22), further revealing why his presentation of a woefully incomplete hypothetical is improper and does not provide data helpful to the factfinder.

Finally, the Likelihood Survey failed to include any context (*e.g.*, a Papaya ad) to orient the respondents and draw the hypothetical closer to a real world scenario for mobile game players. Respondents were just generally told to think about any experience playing a “skill-based” mobile app over the past four years, and as discussed *supra*, Dr. Groehn does not know what game respondents were thinking of, or how they understood “skill-based.”

---

<sup>9</sup> Even if the results of Perception Survey were taken at face value—and they should not be, given the myriad design flaws—a plurality of respondents (nearly 50%) that saw the Papaya video reported that they thought the Papaya game included bot players. (Report ¶ 47, Fig. 10.) So even assuming *arguendo* that Likelihood Survey respondents were thinking about a Papaya game, Dr. Groehn cannot genuinely contend those respondents must have thought they were only playing human opponents.

**3. *Dr. Groehn's Bald Suggestion That Replicating Real World Conditions "Does Not Matter" When Testing "Omissions" Is Meritless***

At deposition, Dr. Groehn again newly attempted to justify his failure to comply with a basic survey standard by once again arguing that he did not need to, and standard survey design principles do not apply, because he was testing an "omission:"

Q: But you omitted everything else from the real-world stimuli from which you drew your mock app store description; right?

A: I don't think that matters.

Q: *You don't think it matters to present consumers advertising as it appears in the real world?* [Objection omitted]

A: *In this context, we are looking at an omission, so it does not matter.*

(*Id.* at 249:10-20; emphasis added.) Once again, Dr. Groehn knew of no authority to support his argument (*id.* at 249:21-250:3), nor would it make any sense that the concerns about mirroring the real world noted above simply disappear in an alleged "omission" scenario; to the contrary, the proper context of an advertisement is essential in order to determine whether information actually has been "omitted" or whether the advertising is actually misleading by "omission."

**C. The Surveys Are Impermissibly Leading and Highly Biased**

Survey questions should be "framed in a clear, precise, and non-leading manner."

*MasterCard Int'l*, 2004 WL 326708, at \*8 (Cote, J.) (citations omitted); *see* Feirman Decl. Ex. 9 (Reference Guide), at 387 (questions should be "clear, precise, and unbiased"). "A survey is not credible if it relies . . . on questions which are inherently suggestive and invite guessing by those who did not get any clear message at all." *Ultreo*, 574 F. Supp. 2d at 352 (citations omitted).

Virtually every element of the Groehn Surveys is highly leading and designed to bias the results in Skillz's favor. Tellingly, the surveys contain none of the standard elements used to minimize bias, such as using open-ended and/or "filter" questions that could reduce guessing and

weed out respondents who do not have a relevant opinion. *See* Feirman Decl. Ex. 10 (ABA Treatise) at 219 (“Filter questions are used to screen out respondents who do not have an opinion on the questions posed by the survey, guessers, and ‘yea-sayers’ who will give an answer that they believe the interviewer is seeking.”) Instead, Dr. Groehn solely provided his closed-ended questions that expressly refer to “bots,” and did not even provide “distractor” questions or response options that could better ensure results are not elevated because respondents were being informed that “bots” were important.

Even beyond this overarching issue, other elements of the surveys are leading and biased. In the Perception Survey, for example:

- Dr. Groehn admits the “mock” description was limited to statements Skillz challenged as deceptive, with all context “stripped” out, leaving solely that challenged language rather than any remotely realistic stimulus;
- Respondents were primed to focus more on the issue of bots in connection with the “mock” description because they viewed that stimulus immediately after having been asked a question *about bots* for the video ad; and
- Respondents were told, prior to watching the video, that they should pay particular attention to the end of the video—running afoul of Dr. Groehn’s testimony that it would *not* be appropriate “to prime or bias respondents by telling them to pay attention to one specific part of the video ad versus another part of the video ad” (Groehn Tr. at 357:15-19).

Additional leading and biasing elements of the Likelihood Survey include that:

- As noted, respondents were baselessly told what they already “thought” about the nature of the opponents in whatever mobile game they had in mind; and
- The question was compound by asking about respondents’ “playing” and “spending your money,” which is problematic both because (i) some respondents may wish to answer affirmatively to only one of those verbs, and (ii) it unnecessarily emphasizes that respondents are putting their own money at stake.

These issues, too, warrant exclusion of the surveys. *See, e.g., Salon FAD.*, 2011 WL 408990-2, at \*9-10 (Cote, J.) (survey questions were “overly suggestive”).

**D. The Surveys Do Not Draw Representative Samples from the Proper Universe**

“[T]he selection of the respondent universe is a ‘crucial step,’ because . . . if the wrong [universe is surveyed], the results are likely to be irrelevant.” *Medisim*, 861 F. Supp. 2d at 179 (citations omitted). “Identification of a survey population must be followed by selection of a sample that accurately represents that population.” (Feirman Decl. Ex. 9 (Reference Guide), at 380.) Dr. Groehn agrees that a sample should “best match the characteristics of the survey population about whom you want to draw conclusions in the real world.” (Groehn Tr. at 337:2-12.) Accordingly, failure to capture the correct sample universe weighs in favor of exclusion. *See, e.g., J.T. Colby*, 2013 WL 1903883, at \*23 (Cote, J.) (excluding survey that, *inter alia*, “failed to test the proper universe of consumers”).

As noted *supra*, Dr. Groehn defined his target population as “those who have used mobile gaming apps” during the past four years. Among other criteria, respondents were required to report that they “played skill-based mobile games in the past four years.” Yet Dr. Groehn’s target population and sample were deeply problematic in at least four ways.

**First**, by only including respondents who already played mobile games, Dr. Groehn excluded *prospective* game players (*i.e.* potential new customers). It is well-established, however, that prospective purchasers are properly included in false advertising and trademark surveys. *See* Feirman Decl. Ex. 10 (ABA Treatise) at 209 (“To test what messages consumers receive from an allegedly false or misleading advertisement, litigants must survey potential prospective purchasers of the advertiser’s product”); Ex. 9 (Reference Guide), at 376. This makes sense because, as Dr. Groehn recognizes, advertising can target new customers “who were not previously aware of a product.” (Groehn Tr. at 345:7-10). Yet despite that recognition, Dr. Groehn only surveyed those with prior experience playing (and thus likely with more knowledge

of) mobile games, ignoring a potentially significant portion of Papaya’s target audience.<sup>10</sup>

**Second**, Dr. Groehn does not understand the composition of his own sample because his screener relied on consumers’ understanding of the undefined term “skill-based.” Respondents only qualified for the surveys if they self-reported that they played “skill-based” mobile games, but Dr. Groehn testified that “skill-based” is a “term of art” with a meaning that “depends,” and he admitted that he has no understanding of how respondents understood that term, let alone whether they shared the same understanding of what a “skill-based” game is. Consequently, there is substantial uncertainty regarding whether and to what extent his survey samples encompass players of Papaya games. And more generally, Dr. Groehn’s reliance on the term without providing any definition means “no meaningful conclusion can be drawn from the respondents’ answers.” *De Lacour v. Colgate-Palmolive Co.*, 16-CV-8364, 2024 WL 36820, at \*5 (S.D.N.Y. Jan. 3, 2024).

**Third**, Dr. Groehn made no attempt to match his sample to demographic characteristics of Papaya’s actual or target consumer base. He did not consider any materials concerning those characteristics, and did not set any quotas to ensure that his sample was a fair approximation of real world consumers. (*Id.* at 343:6-10.) Instead, he relied on his survey vendor to send out survey invitations based on U.S. Census data, and then improperly assumed that response rates to his surveys accurately captured the real-world demographic breakdown.<sup>11</sup> (*Id.* at 337:13-339:2.)

---

<sup>10</sup> Dr. Groehn criticizes Papaya’s experts’ surveys for including prospective Skillz consumers, arguing that all respondents should have prior knowledge of the apps. (*Id.* at 346:16-21.) That is precisely backwards, and Dr. Groehn could not identify any authorities for that proposition. (*Id.* at 348:2-7.)

<sup>11</sup> In his “Supplemental Report,” in an apparent effort to mitigate this failure, Dr. Groehn provided an additional analysis that adjusted the gender split of his respondents to mirror the gender split in Dr. Isaacson’s consumer survey. But Dr. Isaacson tested *Skillz’s* advertising and thus was based on demographic information from Skillz’s document production about *Skillz’s* consumers. Dr. Groehn has no basis to assume that the demographics of Papaya and Skillz consumers are identical. Regardless, Dr. Groehn did not attempt to adjust any other demographic measures (*e.g.*, age).

*Fourth*, Dr. Groehn violated standard survey practices by not knowing the identity or any other information concerning the source of his panels of potential survey respondents. Instead, he relied on his survey vendor to contract with panel providers, and Dr. Groehn exercised no supervision over that process. (*Id.* at 366:2-22.)

**E. Even If the Likelihood Survey Were Methodologically Sound, The Results Are Unhelpful and Prejudicial Because, As Dr. Groehn Admits, They Actually Measure Respondents’ Reactions to Being Told They Were Lied To**

Dr. Groehn contends that the Likelihood Survey “demonstrate[s] that appropriate disclosure as to the presence of bots in [skill-based mobile games] is material to the target audience.” (Report ¶ 70.) But even putting aside the myriad flaws that render it wholly unreliable, the Likelihood Survey is not actually measuring the materiality of bots or advertising statements about bots. Rather, by Dr. Groehn’s own admission, the survey results are capturing a very different concept not pertinent to the materiality of the advertising challenged in this case: how respondents react to being informed that they have been deceived.

This is an inevitable function of the way the question in the Likelihood Survey is framed. Rather than ask respondents how they may be impacted by the presence or absence of bots, or by a particular advertising statement concerning bots, Dr. Groehn instead inquired how likely respondents would be to stop or continue playing and spending their money “*if they were informed* that some of their opponents in the game *whom they thought were real humans* were actually bots.” (*Id.* ¶ 50 (emphasis added).) As noted above, this hypothetical baselessly presumed respondents had particular preexisting “thoughts” about opponents, and told respondents that they were misinformed or affirmatively misled by the app. As a result, it is impossible to know whether the responses were because the respondents actually care about bots, or because they are upset about being told they were lied to.

Dr. Groehn freely acknowledges that his survey results are capturing *both* of the above

concepts. He testified at deposition that he cannot “disaggregate how many people answering [the likelihood question] responded based on their reaction to feeling that they have been deceived versus people just simply learning that there are bots in the game they’re thinking of,” and that he was only capturing the “combination” of those two reactions. (Groehn Tr. at 321:11-325:20.) The results, therefore, are not helpful to the factfinder under Rule 702, and at minimum are prejudicial and confusing such that the survey should be excluded under Rule 403.

**F. The Surveys Suffer From Various Other Methodological Flaws**

Insofar as the Court’s analysis also considers the cumulative effect of all of the flaws in the Groehn Surveys, the following additional issues must be highlighted:

*First*, the introduction to the main questionnaire in the Perception Survey uses the loaded and inaccurate term “wager,” informing respondents they would be shown an ad for an app where you “wager your money to play against other players and the winner of the game is awarded cash or a prize.” (Feirman Decl. Ex. 3 (Questionnaire) at 7.) Despite Dr. Groehn describing “wager” as a “term of art,” he did not define that term for respondents, nor could he explain why he decided to incorporate that gambling-suggestive term in the Perception Survey when he had not used that language in the previously-fielded Likelihood Survey. (Groehn Tr. at 195:12-201:21.)

*Second*, and similarly problematic, is that introduction’s use of the term “winner.” As Dr. Groehn is aware, the use of “winner” in the singular to refer to Papaya’s games is inaccurate, because Papaya offers games where more than one player can win a prize. (*Id.* at 202:9-22.)

*Third*, even putting aside the numerous other missteps in creating his “mock app store description,” Dr. Groehn cherry-picked language for that stimulus from improper sources, including a *Canadian* app store page not targeted at U.S. consumers. (Report at 9 n.9; Groehn Tr. at 274:18-276:17.) And when shown another of the webpages that he listed as a source for

his “mock” stimulus, Dr. Groehn could not locate the language he sourced from that page and suggested that his Report might have included a “typo.” (*Id.* at 286:13-288:20.)

**Fourth**, for reasons never explained in his Report, Dr. Groehn permitted respondents to qualify for his survey even if they responded “I don’t know” to the question asking what mobile devices they have used to play gaming apps in the past four years. (Feirman Decl. Ex. 3 (Questionnaire) at 4.)

**Fifth**, Dr. Groehn included a respondent in the Likelihood Survey who reported that they did *not* understand the questions that had been asked. (Report at ¶ 67, Table 4.) Dr. Groehn inexplicably still considers that respondent “reliable data.” (Groehn Tr. at 395:1-397:8.)

**Sixth**, Dr. Groehn emphasized the importance of his “pretests” before launching his survey, but he did not make changes in response to certain comments from pretest survey takers, and he could not even identify the changes that he did make (including at least one substantive change) because he did not maintain copies of the pretest questionnaires. (*Id.* at 369:22-377:8.)

### **CONCLUSION**

For the many reasons detailed above, the “Perception Survey” and “Likelihood Survey” proffered by Dr. Andreas Groehn are deeply flawed, fundamentally unreliable, and unhelpful to the jury. Those surveys and testimony pertaining thereto should be excluded in their entirety.

Dated: June 27, 2025  
New York, New York

Respectfully submitted,

/s/ Anthony J. Dreyer

Anthony J. Dreyer  
Jordan A. Feirman  
SKADDEN, ARPS, SLATE,  
MEAGHER & FLOM LLP  
One Manhattan West  
New York, New York 10001  
Tel: (212) 735-3000  
Anthony.Dreyer@skadden.com  
Jordan.Feirman@skadden.com

David B. Leland (*pro hac vice*)  
Margaret E. Krawiec (*pro hac vice*)  
Michael A. McIntosh (*pro hac vice*)  
SKADDEN, ARPS, SLATE,  
MEAGHER & FLOM LLP  
1440 New York Avenue, N.W.  
Washington, D.C. 20005  
Tel: (202) 371-7000  
David.Leland@skadden.com  
Margaret.Krawiec@skadden.com  
Michael.McIntosh@skadden.com

*Attorneys for Papaya Gaming, Ltd. and  
Papaya Gaming, Inc.*

**LOCAL RULE 7.1(C) CERTIFICATION**

I, Anthony J. Dreyer, hereby certify that the foregoing memorandum of law complies with the word count limitations set forth in Rule 7.1(c) of the Local Rules of the United States District Court for the Southern District of New York, and contains 8,602 words, exclusive of the caption, table of contents, table of authorities, signature blocks, and this certificate.

Dated: June 27, 2025  
New York, New York

/s/ Anthony J. Dreyer